DOCUMENT RESUME

ED 248 540 CS 208 584

AUTHOR Dvorak, Jack
TITLE Rating the Raters: Some Characteristics of Quill and Scroll's Newspaper and Newsmagazine Judges.
PUB DATE Jan 84
NOTE 20p.; Paper presented at the Mid-winter Meeting of the Association for Education in Journalism and Mass Communication (Baton Rouge, LA, January 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Criteria; Faculty Advisers; *Individual Characteristics; *Journalism Education; *Judges; *Media Research; School Newspapers; *School Publications; Secondary Education
IDENTIFIERS Newsmagazines; *Quill and Scroll

ABSTRACT

To prepare a profile of Quill and Scroll's newspaper and newsmagazine judges, questionnaires were sent to all 57 judges involved in the 1982 competition. Analysis of the 39 responses indicated that the typical judge (1) is female and has judged for two or more years, (2) has taught high school for five or more years, (3) holds certification to teach journalism, (4) has earned 30 or more semester hours of journalism credit, (5) moderately disagrees that knowing a newspaper's previous year's score would help in evaluations, (6) slightly disagrees that there is too much emphasis on graphics and design in the evaluations, (7) agrees that schools with large budgets tend to finish high in contests, (8) agrees that business practices should be part of newspaper evaluations, (9) strongly agrees that a qualified adviser is the single most important factor in producing an award winning publication, and (10) most strongly agrees that she or he is an exacting, thorough, and conservative judge. Other findings showed that the judge knew the previous year's score of each publication evaluated, while statistical comparisons of the years 1980 through 1983 showed no significant differences when previous years' scores were either known or not known. (Author/FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED248540

RATING THE RATERS:

SOME CHARACTERISTICS OF QUILL AND SCROLL'S NEWSPAPER AND NEWSMAGAZINE JUDGES

Jack Dvorak

School of Journalism and Mass Communication

The University of Iowa

A paper submitted for presentation at the mid-winter meeting of the Secondary Education Division of the Association for Education in Journalism and Mass Communication, Louisiana State University, Baton Rouge, Louisiana. January, 1984.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jack Dvorak

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS208 584

ABSTRACT

RATING THE RATERS:
SOME CHARACTERISTICS OF QUILL AND SCROLL'S NEWSPAPER AND NEWSMAGAZINE JUDGES

Judges of the 1982 Quill and Scroll annual newspaper/news-magazine evaluations were sent questionnaires in 1983--39 of 57 (68.42%) responded. The typical judge: 1) is female who has judged newspapers for 2 or more years, 2) is a high school teacher of 5 years or more, 3) holds certification to teach journalism within the state, 4) has earned 30 or more semester hours of journalism, 5) moderately disagrees that knowing the previous year's score would help make decisions about a current newspaper evaluation, 6) slightly disagrees that there is too much emphasis on graphics and design in the evaluations, 7) agrees that schools with large budgets tend to finish high in contests, 8) agrees that business practices should be a part of newspaper evaluations, 9) strongly agrees that a qualified adviser is the single-most important factor in an award-winning publication, and 10) most strongly agrees that she or he is an exacting, thorough and conservative judge.

In years before 1982, judges knew the previous year's score of each newspaper evaluated. Statistical comparisons of the years 1980 through 1983 showed no significant differences when previous years' scores were known (1980, 1981) and when they were not known (1982, 1983).

Judges' perceptions of whether or not knowing previous years' scores influenced their decisions in making current newspaper evaluations were crosstabulated with various demographic and psychographic items on the questionnaire. Several of the statistically significant relationships are described.

Rating the Raters: Some Characteristics of Quill and Scroll's Newspaper and Newsmagazine Judges

Ours is a competitive society. As such, even the most noble of human endeavors is subjected to some type of recognition system following elaborate schemes of evaluation.

Nobel prizes in various categories are presented regularly to the world's leaders. In journalism, various organizations present awards for achievement in reporting, public service and the like.

While awards like the Nobel Peace Prize or the Pulitzer Prize in journalism are among the most prestigious, at times the awards' panels are subjected to scrutiny by their respective constituent groups. For example, the Pulitzer prizes have been tarnished a bit of late after the Janet Cooke scandal; also, journalistic leaders have questioned the Pulitzer judging panel, both before and after the Janet Cooke episode, regarding the makeup of the panel, the politics of the awards and other procedural guidelines.

High school journalism also has its coveted national awards and rating services that issue those awards. This paper examines one such organization--Quill and Scroll--and its personnel who judge the annual newspaper and newsmagazine contests.

While this organization (as well as others such as the Columbia Scholastic Press Association and the National Scholastic Press Association) has been the subject of other studies, no study--at

least not since 1980--has been found that examines the judges' perceptions in making individual evaluations of newspapers.

Just as criticism has been levied against the prestigious national professional award systems, dissatisfaction with scholastic rating services is easily found.

H.L. Hall collected data from 216 high school advisers in 1981 in which their concerns related to rating services were aired.[1]

A comparison of costs was done by Rasmussen in 1981.[2] Her study examined various positive and negative aspects of several of the national rating services, especially as they related to the price to the school media being evaluated.

Other studies have attempted to evaluate how the major national rating services compare to each other when schools send similar issues of newspapers to the same national services (in this case CSPA, NSPA and Quill and Scroll).[3]

In 1980 Blick studied various factors significant in schools' achieving high ratings in CSPA evaluations.[4]

Prentice examined 5 ratings' services--including Quill and Scroll, NSPA, CSPA, the American Scholastic Press Association and the Texas Interscholastic League Press Conference--and determined that in four of them too strong an emphasis on graphics and design was included in the point systems.[5]

In another article, Prentice explored various internal and external criteria used by the services in their evaluative schema. He found substantial flaws in four of the five systems.[6]

While these studies are helpful in making an overall assessment

of rating services and what some of the clients of these services find to be positive and negative about them, the examination of judges themselves has been lacking.

Methodology

All 57 Quill and Scroll judges--33 women (58%) and 24 men (42%)--who participated in newspaper/newsmagazine evaluations in 1982 were sent a one-page questionnaire in the spring of 1983. Various aspects of judging on both affective and cognitive levels were included, as were several demographic items.

Thirty-nine questionnaires were returned for a 68.42% response rate. Besides examination of frequencies, data were subjected to various crosstabulations.

In addition to that data, this study also examines comparative scores for schools involved in Quill and Scroll evaluations from 1980 through 1983. Only those schools that had the same newspaper adviser throughout that period were used in the examination because it is felt that the adviser assures a level of stability to a publication; thus, a comparison of these scores using various statistical measures has more credibility--even though in almost every case the newspaper was judged each of the years by a different person since it is Quill and Scroll's policy to do this.

Specifically, scores of those years were examined because it had been a long-standing policy of Quill and Scroll to include the previous year's score to a judge before that person began evaluating the publication. This practice was discontinued starting with the 1982 contest. The Pearson product-moment correlation was set up with

the years 1980, 1981, 1982 and 1983 included to see if real differences in scores existed. Also, t tests were run comparing 1981 with 1982 scores--along with 1980/1981 and 1980/1982--to see if withholding previous years' scores made differences in outcomes.

Results and Discussion

Profile of a Quill and Scroll Judge. Using results of the questionnaire, one is able to examine a typical judge--or at least a composite of what one might consider typical--before a deeper look at the findings.

The Quill and Scroll judge is usually a woman who has judged for 2 or more years. She spends 3 to 4 hours judging one newspaper, although chances are more likely than not that she spends more than 4 hours per entry. In addition to Quill and Scroll, she judges for 2 or more other press rating services, although there is a strong likelihood that if she does not judge for 2 or more others she judges for no other organization.

This typical judge is a high school teacher with 5 years or more in her current job. She teaches in a state in which she holds certification to teach journalism. During undergraduate and graduate study, she earned 30 or more semester hours of journalism.

Table 1 shows the results (N=39) of Quill and Scroll judges' perceptions of the ideal point breakdown in various categories of the Score Book.[7] Average scores might be considered the viewpoint of the "typical" judge. Note that actual scores of the Score Book are similar. This might be attributed to the judges' familiarity with the booklet itself.

TABLE 1

Comparison of Ideal vs. Actual Point Allocation
(N=39)

Item	Ideal Allocation	Actual Allocation	Difference
Policy Guidelines	12.64%	18%	-5.36
Coverage	21.10%	20%	+1.10
Writing and Editing	32.36%	35%	-2.64
Display and Design	22.39%	15%	+7.39
Business Practices	12.28%	12%	+0.28

The area of greatest difference between actual booklet allocations and the judges' ideal is "Display and Design," in which judges felt 22.39% of the point total should be granted while Quill and Scroll currently allows only 15% of its allocation to display/design.

The second highest area of disagreement on point allocation falls in the "Policy Guidelines" category. Here, judges indicated that only 12.64% should be granted while Quill and Scroll currently allocates 18%.

The three other areas of coverage, writing and editing, and business practices show negligible differences between judges' perceptions of the ideal and actual point allocations in the Quill and Scroll Score Book.

Table 2 shows the assessment of judges of some statements thought to be either controversial or crucial to the judging process. The statements were derived from one of the key questions in the current study--whether or not knowing previous year's scores influenced judges' scoring in the present year--and also derived from some current

literature regarding the structure of national rating services.

Regarding the questions related to knowing previous years' scores, the typical judge disagreed moderately that knowing those scores would help in making decisions about this year's newspapers. A related question showed a neutral response: that if last year's score were known, it would tend to influence this year's final score. However, it should be noted that 17 of the respondents either disagreed strongly or disagreed with the statement while 15 of the judges agreed strongly or agreed with the statement. Only 7 checked the "neutral" or "not sure" response. Thus, there was a wide distribution of opinion on this statement even though numerical average was neutral.

Slight disagreement was also registered in the area of graphics. Judges indicated a negative reaction to the statement that there is too much emphasis on graphics and design, with basics of writing, reporting and editing suffering. While the average is just under 3.0 (neutral), it should be pointed out that 12 judges disagreed and 3 strongly disagreed, while only 9 indicated any sort of agreement. Fifteen judges checked the "neutral or not sure" category.

General agreement was registered in Table 2 with the statement that schools with the biggest budgets for newspapers tend to have the best looking newspapers, and thus usually finish high in contests. (A study by Blick indicates a positive relationship between amount of funding and the winning of awards.[8])

Stronger agreement was registered on item F., even though it was posed in the negative. The 2.41 score indicates a negative result to

the statement "business practices should not be included in a newspaper's evaluation" (emphasis mine). Thus, there is general agreement that business should be a part of the score booklet for newspapers.

TABLE 2

Judges' Perceptions of Conditions Related to Newspaper Evaluation (N=39)

Rating*	Statement
2.87	A. Knowing final scores from the previous year's Score Book would help me make decisions about this year's newspapers.
3.00	B. If I know the score from last year's evaluation, it would tend to influence this year's final score.
2.82	C. In high school newspapers today, there is too much emphasis on graphics and design, with basics of writing, reporting and editing suffering.
3.26	D. Schools with the biggest budgets for newspapers tend to have the best looking newspapers, and these usually finish high in contests.
4.08	E. I tend to be an exacting, thorough and conservative judge.
2.41	F. Business practices should not be included in a newspaper's evaluation.
3.82	G. A qualified adviser is the single-most important factor in an award-winning school newspaper.

*Based on 5-point scale in which "5"=strongly agree, "4"=agree, "3"=neutral or not sure, "2"=disagree and "1"=strongly disagree

---

Second highest agreement in Table 2 was to the statement that a qualified adviser is the single-most important factor in an award-winning school newspaper. Only six respondents disagreed with the statement, six indicated neutral or not sure, and 27 agreed (with 12 of those strongly agreeing).

Strongest agreement of judges--and the final statement summarizing the "typical" Quill and Scroll judge, was to a statement about their self-perceptions in roles as judges. Statement E. in Table 2 shows general agreement: "I tend to be an exacting, thorough and conservative judge." Unfortunately, "conservative" muddled the intent of the researcher a bit, as indicated by a few respondents who wrote back that they agreed with "exacting" and "thorough" in the statement but not the third and final part. This is due, no doubt, to the various connotations of the word "conservative" as related to other social and political ideologies. However, the intent was to see if judges saw themselves as setting high standards for the awarding of points in contests rather than "liberally" awarding points in cases where doubt existed or where work not in keeping with generally accepted newspaper principles might be given more points than deserved.

Some of these statements will be examined further through crosstabulations later in the paper.

Effect of Knowing Scores. Through the Quill and Scroll newspaper and newsmagazine evaluations of 1981, and having been a practice for at least a dozen years before that, judges were sent previous year's scores with each publication to be judged. While this undoubtedly had several advantages, especially in schools that had the same newspaper adviser, it also seemed to have the disadvantage of unnecessarily influencing a judge to evaluate so that no major deviations from the previous year occurred. The practice was discontinued beginning with the 1982 evaluations.

In order to test for differences in recent years, two tests were completed. Both t test of matched pairs and the Pearson product-moment correlation would indicate that no statistically significant changes occurred between the years when judges knew previous year's scores (1980, 1981) and when they did not (1982, 1983).

Table 3 illustrates the results of t tests in which three comparisons were drawn: 1) 1982 with 1981, the principal comparison of interest in the current study; 2) 1980 with 1981, comparing differences between years when judges knew the previous year's scores; and 3) 1980 with 1982, comparing a year when scores were known with the first year in which they were not.

These comparisons are based on 99 schools that participated and which had the same newspaper adviser during the years 1980-1982. In analyzing the data, one notes that aside from no significant differences being recorded, standard deviations and mean scores for the three years do not differ enough to make any general conclusions. For example, in 1981 the mean score of the 99 newspapers in the evaluations was 894.99, while in 1982 it was 896.85--a difference of only 1.86.

One might reasonably expect the score to be a bit higher when the previous year's score is known. It was higher when the previous year's score was unknown.

Similarly, 1980, the first year a new scorebooklet was used (revised edition), the scores were lower than during the second year the booklet was used and the previous year's scores were known (1980/1981). Of the three years, the standard deviation was

highest during 1980 even though the mean was lowest--again, possibly due to the unfamiliarity of judges with the revised Score Book.

TABLE 3

T-Test Comparisons of 1982/1981, 1980/1981, and 1980/1982 (N=99)

Year	Mean	Standard Deviation	T Value	df	p*
1982	896.85	54.49	0.41	98	0.68
1981	894.99	52.36			
1980	887.46	65.19	-1.57	98	0.12
1981	894.99	52.36			
1980	887.46	65.19	-1.67	98	0.10
1982	896.85	54.49			

*No 2-tailed probabilities were significant

During the four years the revised Score Book had been used, 1980-1983, final scores from 68 school newspapers that retained the same adviser all four years were correlated using Pearson product-moment correlations (Table 4). The lowest mean score occurred when previous year's scores were known by judges--1980 (896.16). Standard deviation that year was 66.19--highest of the four years analyzed in this four-year grouping of contestants. Other means and standard deviations of the years tested include 904.38 (s.d.=49.66) in 1981; 903.60 (s.d.=52.84) in 1982 (the first year scores of previous years were not provided); and 900.19 (s.d.=57.02) in 1983.

Using Guilford's interpretations of coefficient probabilities,[9] it would be safe to conclude that all four years correlated in

Table 4 show moderate correlations with substantial relationships--except the coefficient between 1981 and 1982, which shows a high correlation coefficient (.72) and a marked relationship.

TABLE 4

Correlation Coefficients of Final Scores, 1980-1983
(N=68)

	1980	1981	1982	1983
1980	****	.658 p<.001	.519 p<.001	.472 p<.001
1981	.658 p<.001	****	.725 p<.001	.682 p<.001
1982	.519 p<.001	.725 p<.001	****	.587 p<.001
1983	.472 p<.001	.682 p<.001	.587 p<.001	****

This highest relationship occurs between the two key years of this part of the study: 1981, when scores were known, and 1982, the first year judges were not given the previous year's score.

The high correlation between 1981 and 1982 (as well as those among the various 4-year combinations) further indicates the lack of influence judges' knowledge of previous years' scores had on the judging process. In all years that were correlated, substantial and significant relationships are apparent no matter what the mix of judges' knowledge or lack of knowledge of previous years' scores. This indicates that it is highly improbable results of test scores changed from one year to the next due to knowledge or lack of knowledge of previous years' scores.

Other Relationships. By using the chi-square statistical test, other insights into the data on the questionnaire of judges are possible. A significant chi-square statistic indicates the lack of goodness of fit, and from this it is possible to examine various trends and relationships among Quill and Scroll judges that might not otherwise be apparent from a simple analyzation of the questionnaire.

Years as Judge. The number of years a judge was involved with Quill and Scroll significantly related ($p<.04$) with whether or not the judge felt knowing final scores from the previous year would influence her/him in making decisions about the current year's newspaper scores.

TABLE 5

Crosstabulation of Years Judged by Knowing Final Scores

Years as Judge	Knowing Final Scores			
	Disagree	Neutral	Agree	
1	7	1	1	(9)
2-4	5	5	5	(15)
5+	5	1	9	(15)
	(17)	(7)	(15)	39

$X^2 = 10.076$ df = 4 $p < .04$

Interestingly, those who have judged for 5 years or longer tended to agree that knowing final scores would help them in making decisions while those who had judged for only 1 year tended to disagree. Those who judged between 2 and 4 years were split on that question with 5 agreeing, 5 neutral or not sure and 5 disagreeing.

Similarly, a significant relationship ($p < .01$) is the result of the crosstabulation between the number of years judged and whether or not that judge would actually be influenced by knowledge of the previous year's score (Table 6).

TABLE 6

Crosstabulation of Years as Judge by Influenced by Scores

Years as Judge	Influenced by Scores?			
	Disagree	Neutral	Agree	
1	1	1	7	(9)
2-4	5	5	5	(15)
5+	11	1	3	(15)
	(17)	(7)	(15)	39

$X^2 = 13.909$ df = 4 $p < .01$

This bitof data actually explains the relationship examined in Table 5. It appears that those who have judged only 1 year agree that they would be influenced by knowledge of the final score from last year, and thus they more readily agree that they should not know the score. Those who have 5 or more years of judging experience, meanwhile, disagree that they would be influenced. (This could explain why that same group in Table 5 indicated that knowledge of the previous score would help them in making decisions. That is, they feel they could be a better judge by knowing the score while at the same time they feel they would not be influenced by that knowledge.)

Journalism Certification. It was also hypothesized that those judges with journalism certification might have significant relationships

with other variables. This is true for one relationship (Table 7).

TABLE 7

Crosstabulation of Journalism Certification with Budgets

Journalism Certification	Big Budgets=Big Scores?			
	Disagree	Neutral	Agree	
Yes	11	3	11	(25)
No	1	4	9	(14)
	(12)	(7)	(20)	39

$X^2 = 6.055$ df = 2 $p < .05$

Those who are certified to teach journalism are about evenly split on the issue of whether or not those schools with big budgets for school newspapers are also most likely to finish high in contests ($p < .05$). However, those without journalism certification--including several judges from the professional journalism ranks who don't need certification--are more likely to agree that big newspaper budgets help newspapers attain high final scores.

Semester Hours in Journalism. A corollary hypothesis was that significant relationships would result when examining the cross-tabulation of the number of hours a judge had in journalism and the person's attitude about knowing the previous year's score.

As can be seen in Table 8, judges with 11 or fewer hours of college journalism training--those who tend to be uncertified--are about evenly split on the issue "knowing final scores from the previous year's Score Book would help me make decisions about this year's newspapers." Those in the mid-range of formal journalism

training who have had between 21 and 29 semester hours most strongly disagree with that statement. Those with 30 or more semester hours more strongly agree than disagree with the statement with 12 agreeing, 8 disagreeing and 3 indicating neutrality or uncertainty.

TABLE 8

Crosstabulation of # of Semester Hours by Knowing Final Score

Semester Hours in Journalism	Knowing Final Scores			
	Disagree	Neutral	Agree	
6-11	3	4	2	(9)
21-29	6	0	1	(7)
30+	8	3	12	(23)
	(17)	(7)	(15)	39

$X^2 = 11.366$ df = 4 $p < .03$

Interestingly, no significant difference was detected when crosstabulations were run comparing number of semester hours judges had taken in journalism and the statement "if I know the score from last year's evaluation, it would tend to influence this year's final score."

Demographic Comparisons. In order to examine characteristics of judges more closely, some final crosstabulations of the demographic data were tabulated. Only significant relationships will be reported.

Those who spend 3 or more hours judging a newspaper are more likely to have journalism certification than those who do not ($p < .02$). For example, of those judges who spent less than 3 hours per newspaper judged, only 5 had journalism certification while 9 did not. However,

of those who spent between 3 and 4 hours 11 of 12 had certification. Of those judges who spent more than 4 hours judging each paper, 9 had certification while only 4 did not.

Of significance in the same consideration--hours spent judging each entry--is the "current profession" category. Here, it can be seen that those who spent less than 3 hours judging each category are generally from the non-high school teaching ranks ($p < .001$). Only 5 of 28 high school teachers spent less than 3 hours judging each entry while fully 9 of 14 who checked the less than 3 hours category were not high school teachers. Conversely, 11 of 12 respondents who spent 3 to 4 hours judging were high school teachers, and 12 of 13 who spent more than 4 hours per entry were high school teachers. In the current study, then, 11 of 39 judges were not high school teachers, and of those 11 only 2 spent more than 3 hours per entry.

Of the judges in the current study, those with 5 years or more in their current jobs more likely had journalism certification than those in present positions for 4 or fewer years ($p < .03$). Only 2 of 7 judges with less than 5 years in present jobs had certification while 23 of 32 judges with 5 years or more in present jobs had journalism certification within their states.

Also significant ($p < .01$) was the relationship between the number of years in a judge's present position and her or his current profession. Only 2 of 7 judges who were in present positions for 4 years or less were high school teachers while 26 of 32 who were in present jobs 5 or more years were high school teachers.

Reference Notes

[1]H.L. Hall, "Turnabout is Fair Play," Communication: Journalism Education Today 15 (Winter, 1981), p. 7.

[2]Pat Rasmussen, "It's Not the Cheapest Advice," Communication: Journalism Education Today 15 (Winter, 1981), p. 2.

[3]Nancy Davis, "Evaluating the Evaluators: A Comparative Study of High School Newspaper Critique Services," (Paper presented to the Secondary Education Division of the Association for Education in Journalism Annual Convention, East Lansing, Michigan, August 1981), p. 2.

[4]Thomas Edward Blick Jr., "High School Newspapers: Factors Significant in Achieving High Ratings," (A Master's Paper in Journalism, The Pennsylvania State University, July 1980), p. 6.

[5]Tom Prentice, "It's Time to Correct the Graphics Skew in Rating Services," Communication: Journalism Education Today 17 (Winter, 1983), p. 14.

[6]Tom Prentice, "A Pedagogical Approach to Rating and Evaluating Student Publications," Communication: Journalism Education Today 17 (Winter, 1983), p. 12.

[7]For comparative purposes, see articles by Davis and Prentice listed above.

[8]Blick, p. 6.

[9]J.P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill, Inc., 1956), p. 145.